



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2011

A proposal and prototype for a Research Risk Repository to improve the protection of research participants

Rid, A ; Wendler, D

Abstract: A Research Risk Repository has the potential to significantly improve the consistency and accuracy of the evaluation of research risks. However, currently available data are generally insufficient for this purpose. Future collection and maintenance of data on the risks posed by research interventions will help to ensure that research participants receive appropriate protection and promote important research consistent with adequate subject protection.

DOI: <https://doi.org/10.1177/1740774511414595>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-54889>

Journal Article

Accepted Version

Originally published at:

Rid, A; Wendler, D (2011). A proposal and prototype for a Research Risk Repository to improve the protection of research participants. *Clinical Trials*, 8(6):705-715.

DOI: <https://doi.org/10.1177/1740774511414595>

A Proposal and Prototype for a Research Risk Repository to Improve the Protection of Research Participants

Annette Rid, MD^{1,2}
David Wendler, PhD¹

¹Department of Bioethics
NIH Clinical Center,
Bethesda, MD, United States

²Institute of Biomedical Ethics
University of Zurich,
Zurich, Switzerland

NOTE: This is a version of the article that was accepted for publication in *Clinical Trials*. It does not incorporate the changes made by the journal after acceptance.

Running Head: Research Risk Repository

Total Words: 4128 words

Author for Correspondence:

Annette Rid, MD
Institut für Biomedizinische Ethik
Universität Zürich
Pestalozzistrasse 24
CH-8032 Zürich
+41 44 634 83 68 (p)
+41 44 634 83 89 (f)
annette.rid@gmail.com

Grant support: Swiss National Science Foundation (AR)

Background Accurate and consistent risk assessment is vital in biomedical research.

Yet, current evaluation of the risks of research interventions often does not take into account the relevant empirical data. This approach raises concern that current practice may not be protecting research subjects adequately, or that it may be blocking acceptable research.

Purpose To propose and evaluate the possibility of creating and maintaining a Research Risk Repository which would make empirical data on the risks of research interventions available to IRBs, investigators, funders, and others.

Methods Analysis of the usefulness of a Research Risk Repository and evaluation of whether currently available empirical data are sufficient to establish such a repository.

Results Creation of a Research Risk Repository would provide a vital resource for systematically and accurately evaluating the risks of clinical research. Realizing this goal requires data that have at least 4 characteristics: 1) **trustworthy**: to ensure credibility to all stakeholders; 2) **robust**: to support confident risk determinations; 3) **inclusive**: to cover all potential harms of the interventions under review; and 4) **comprehensive**: to determine which factors influence the risks of the interventions under review. Evaluation of existing data reveals that they satisfy these requirements for only a few research interventions and, even in those cases, only to a limited extent. Gaps in the currently available evidence highlight the need for systematic collection and maintenance of data on the risks posed by research interventions.

Limitations Creation and maintenance of a Research Risk Repository would be laborious and costly and require regular updating as new data are collected, and as new practices and interventions emerge.

Conclusions A Research Risk Repository has the potential to significantly improve the consistency and accuracy of the evaluation of research risks. However, currently available data are generally insufficient for this purpose. Future collection and maintenance of data on the risks posed by research interventions will help to ensure that research subjects receive appropriate protection.

318 words

Key Words: Risks, Harms, Clinical research, Research subjects, Subject Protection

THE ETHICS OF RESEARCH RISK AND EMPIRICAL DATA

The ethical appropriateness of biomedical research depends on accurate and consistent assessment of the risks faced by research subjects. Accurate and consistent assessment of the risks of biomedical research in turn requires access to empirical data on the risks posed by research interventions. Yet, few systematic data are available on the risks of even common research interventions. As a result, institutional review boards (IRBs) and others currently evaluate the risks of biomedical research based largely on intuition alone.

The problems with relying on intuition alone are highlighted by extensive empirical data from psychology, which reveal that intuitive evaluation of risks is strongly influenced by numerous cognitive biases (1-4). For example, people tend to judge familiar activities as less risky than unfamiliar ones. This bias increases the chances that those familiar with an intervention will judge it to be low risk, while those not familiar with the intervention will judge it to be higher risk.

This concern is reinforced by empirical studies which find significant variation in the assessment of research risks (5-7). In a survey of 188 IRB chairpersons in the U.S., 23% categorized allergy skin testing as minimal risk in healthy 11-year olds, 43% categorized the same procedure, in the same population, as a minor increase over minimal risk, and 27% categorized it as more than a minor increase over minimal risk (5).

It is unlikely that the actual risks of allergy skin testing in healthy 11-year olds vary to this extent between sites, suggesting that the current evaluation of research risks does not always reflect the risks subjects face. Some may be underestimating the risks posed by research interventions, and thereby failing to protect participants from excessive risks. Others may be overestimating the risks, and thereby inadvertently blocking acceptable

research. Variation in risk judgments across sites also poses a significant obstacle to important research that must be conducted in multiple sites, such as research with children and research on rare diseases.

The present paper proposes and evaluates the feasibility of establishing a *Research Risk Repository*—a centralized database that would store and make available to IRBs, investigators, funders, and others, systematic data on the risks of research interventions. Creation of a Research Risk Repository would provide the data necessary for making accurate and consistent risk evaluations, thus helping to better protect research participants from excessive risks, while promoting important research consistent with adequate subject protection.

THE NEED FOR A RESEARCH RISK REPOSITORY

To assess the risks of research interventions, such as lumbar puncture and allergy skin testing, IRBs could rely on empirical data obtained at their local institutions. This approach has the virtue of basing risk assessments on data that reflect local circumstances. However, it leaves IRBs without a way to systematically evaluate the risks of interventions performed at their institutions for the first time. Furthermore, while local risk data often are necessary for judging research risks accurately, they typically are not sufficient. Data from one or even several institutions often will not include enough data points to capture low probability events, such as the risk of anaphylactic shock from allergy skin testing. To capture these events, and adequately protect research subjects, IRBs need access to large datasets on the risks of research interventions.

One way to address this need would be to establish a *Research Risk Repository*—a database that maintains the available data on the outcomes of research interventions, and makes these data widely accessible. A Research Risk Repository would help IRBs to assess accurately the risks of research interventions. IRBs also could use the repository to identify ways to reduce risks to participants. For example, if the repository reveals that using a narrower needle significantly reduces the likelihood of postdural headache following lumbar puncture, reviewers could consider requiring investigators to use narrower needles.

Investigators also could rely on the repository when writing consent forms and discussing research risks with potential subjects. Similarly, the dataset would allow quality managers to compare the outcomes for research interventions across institutions, and use this information to improve institutional practices. Finally, investigators, sponsors, and IRBs could use the repository to identify gaps in the available data and consider ways to collect additional data. For example, IRBs could require collection of more data as part of the study when the existing evidence suggests that the risks of the interventions being used are acceptable, but the available data are incomplete. This approach would allow IRBs to confirm the assumption on which they approve the study – namely, that the risks are acceptable. In addition, funders, such as the NIH, could identify gaps in the empirical data and invite investigators to collect the needed data.

FOUR CRITERIA

To provide the data needed to evaluate the risks of research accurately, a Research Risk Repository would need to satisfy at least four criteria.

The data in the repository should be trustworthy to the various stakeholders in research.

Perhaps the best way to promote trust would be for the repository to be operated by an independent non-profit organization. The work of this organization could be overseen by a committee of experts. A searchable online version of the repository would help to increase accessibility and promote accountability and trustworthiness. To avoid discouraging institutions from contributing their data to the repository, the online version could anonymize the origins of the data.

The evidence should be sufficiently robust and up to date to support confident risk determinations. The data in the repository should capture as many observations as possible, made in diverse settings. The available data also should be updated as new data are collected. This is particularly important when techniques change in ways that likely influence their risks, such as the introduction of narrower needles for lumbar puncture or CT scanning techniques using increased levels of radiation.

The data in the repository should include the types and magnitudes of all relevant potential harms. To protect research subjects from excessive risks, IRBs should evaluate all the potential physical, psychological, social, and economic harms posed by research interventions. This view differs from some commentators who argue that potential negligible harms, such as the chance of local itching as the result of an allergy skin test, constitute mere inconveniences – not risks – and thus can be excluded from risk evaluations (8). However, potential harms that are negligible on their own can have

additive or synergistic effects. To address this possibility, a Research Risk Repository should include data on the types and magnitudes of all potential research harms.

The data should include information on which factors influence the risks of research interventions. The risks posed by some interventions vary depending on the circumstances: how the intervention is performed (e.g., the risk of postdural puncture headache after lumbar puncture increases with larger needles) or who undergoes it (e.g., the risks of sedation increase with underlying disease). IRBs should base their risk determinations on data relevant to the research under review. A Research Risk Repository could facilitate these determinations by stratifying the data by relevant contextual factors, tailored to individual interventions. For example, the repository could classify the needle size used for lumbar puncture as small (24-27 G), medium (20-22 G), or large (16-19 G). Similarly, where relevant, it will be important to specify the extent to which interventions pose greater or lower risks in specific populations, such as children.

EVALUATION OF FEASIBILITY

Review of Existing Risk Data

The most straightforward way to create a Research Risk Repository would be simply to collect and collate existing empirical data on the risks of research interventions. To maximize its usefulness, the repository should focus, at least initially, on common research interventions that are used in contexts which do not offer participants the potential for clinical benefit, such as biopsies or blood draws performed for purely research purposes. Research interventions that pose risks to subjects without a

compensating potential for clinical benefit raise the greatest ethical concern. Moreover, whether the risks of these “non-beneficial” interventions are classified as “minimal” or greater than “minimal” has important practical implications. For example, most current guidelines and regulations allow “non-beneficial” interventions to be performed in healthy children only when they pose minimal risk.

To evaluate whether existing data satisfy the four requirements of a Research Risk Repository, we conducted a preliminary literature search on the risks of common research interventions (figure). To be included in the search, the interventions had to meet the following criteria: 1) performed routinely in the research setting; 2) often performed for research purposes only (and sometimes controversial in this “non-therapeutic” context); and 3) sufficient data are likely available on the risks posed by the intervention.

We started by classifying research interventions in 10 categories: 1) asking questions (e.g., survey of health-related behaviors); 2) physical examination (e.g. blood pressure measurement); 3) imaging (e.g., ultrasound, CT scan); 4) measuring electrical activity (e.g., ECG, EEG); 5) sampling fluids or cells (e.g., blood draw, apheresis); 6) removing tissue (e.g. skin or muscle biopsy); 7) instilling substances (e.g. allergy skin testing); 8) genetic testing (e.g. genetic test on previously obtained sample); 9) hospitalization (e.g. overnight stay in hospital); 10) record keeping (e.g. record confidential medical information). In consultation with 25-30 clinical or biomedical researchers, IRB members, and experts in research ethics, we drafted a list of 33 research interventions that met the first two inclusion criteria (table 1). We then performed brief literature searches to determine whether the existing data were likely to satisfy the third inclusion criterion of including sufficient data points on the risks of the 33 interventions.

This process, together with consultation with our group of experts, led to the exclusion of sixteen interventions for which sufficient data were unlikely to be available. For example, we could not identify pertinent data on repeat blood sampling, defined as a series of 5-10 and 10-20 blood draws of a maximum of 10 cc each. Although some studies record the volume of diagnostic blood loss in the inpatient setting, they often do not estimate the clinical impact. Analogies to blood donation are weak because the data on the risk of anemia and iron depletion in this setting are scarce and/or outdated (9). Similarly, the clinical implications of exposure to low dose radiation remain unclear (10). We thus excluded imaging interventions involving radiation, such as X-rays and CT scans.

To identify the highest quality datasets for the remaining 17 interventions (see appendix 1), we retrieved citations from PubMed combining the MeSH term for the given research intervention (e.g. “skin tests” for allergy skin test) with MeSH terms for risk (e.g. risk OR adverse effects), and included all studies published in English on or before 10 January 2009 (see table 2 for search strategy). The first author (AR) then contacted an independent clinical expert/researcher with the results. One expert per procedure was contacted. Only two of the consulted experts worked outside of NIH at the time. The experts helped to select the highest quality references and identified important papers or book chapters that were not retrieved by the search, often in consultation with their colleagues.

The first author (AR) or a research assistant catalogued the retrieved risk data for each intervention in a table (see table 3 for the template table), arranging potential physical harms and associated likelihoods (given per 100,000) according to body systems.

Potential psychological and social or economic harms were listed separately, if available. Data that included pediatric populations were shaded in grey. Each table also indicates the total number of observations in the cited studies (i.e. the total number of interventions performed or the number of observed cases in the context of case reports), and provides relevant information about the respective study contexts. A 95% Confidence Interval (95% CI) for proportions was calculated using the Wilson formula (“Wilson score interval” (11)). The analysis was performed with open source software of the “R foundation for statistical computing” (<http://www.r-project.org/>).

The identified clinical experts/researchers provided the first author (AR) with a detailed description of the potential harms resulting from the given research intervention, including the type, duration, and reversibility of each potential harm. The experts also provided likelihood estimates for potential harms not provided in the literature, as well as general comments on the quality of the data obtained and relevant background information on the given intervention (online appendix 1). Finally, the experts helped to develop “summary tables” for each intervention (online appendix 2). With the exception of the table on moderate sedation for children, the summary tables do not reflect data collected in pediatric populations.

Figure. Process used to test the feasibility of creating a *Research Risk Repository*.

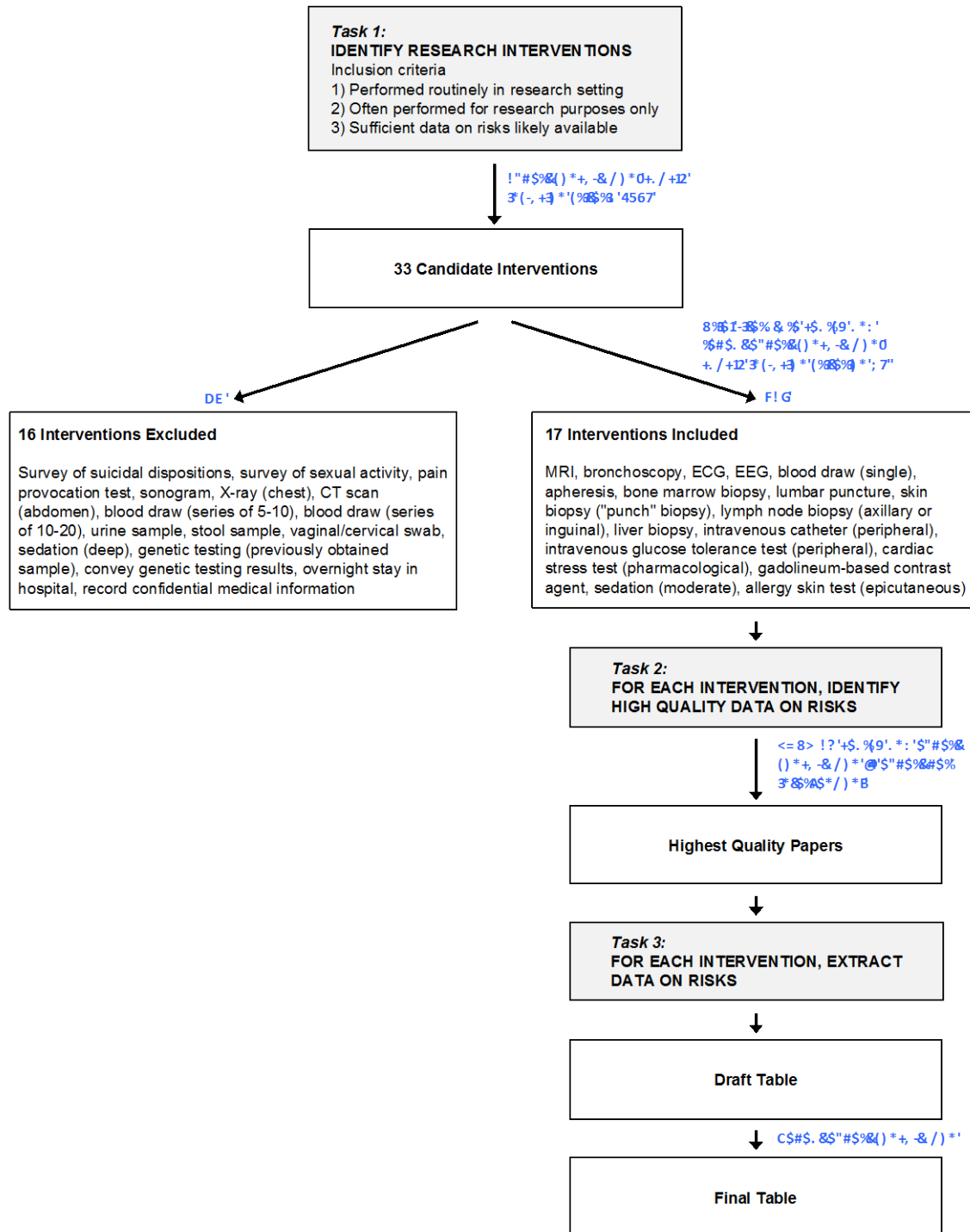


Table 1. Research interventions included in the preliminary review of the existing data (listed in alphabetical order). The tables listing the data of each intervention, as well as summary tables developed in consultation with experts, are available as in the online appendix. With the exception of the table for moderate sedation in children, the summary tables do not reflect data collected in pediatric populations.

Allergy skin test (epicutaneous)

Apheresis

Blood draw (single)

Bone marrow biopsy

Bronchoscopy

Cardiac stress test (pharmacological)

Electrocardiogram (ECG)

Electroencephalogram (EEG)

Gadolinium-based contrast agent

Intravenous catheter (peripheral)

Intravenous glucose tolerance test (peripheral)

Liver biopsy (percutaneous)

Lumbar puncture

Lymph node biopsy (axillary or inguinal)

Magnetic Resonance Imaging (MRI)

Sedation (moderate, i.e. with purposeful response to verbal or tactile stimulation)

Skin biopsy (“punch” biopsy)

Table 2. PubMed Search Strategy.

MeSH term for research interventions (e.g. “skin tests” for allergy skin test)

AND

Risk [MeSH] or adverse effects [MeSH] or mortality [MeSH] or incidental findings

[MeSH] or (adverse effects AND psychology) [MeSH] or stress, psychological

[MeSH] or adaptation, psychological [MeSH] or socioeconomic factors [MeSH]

or stereotyping [MeSH]

AND

English only

Table 3. Template table used to extract the available risk data. Rows reporting data from pediatric populations were shaded in grey. Notes on the quality of the data and background information about the procedure, as well as the list of references, were recorded at the bottom of each table.

[Name of Intervention]: Risks and Complications

Harm	Likelihood (per 100,000)	95% CI (per 100,000)	N	Study Context	Ref.
1. Physical Harm					
a. Morbidity					
General symptoms:					
E.g. Malaise					
Neurological System:					
E.g. Headache					
Cardiovascular System:					
E.g. Syncope					
Respiratory System:					
E.g. Wheezing					
Gastro-intestinal System:					
E.g. Nausea and Vomiting					
Kidney and urinary tract:					
E.g. Cystitis					
Reproductive System:					
E.g. Prostatitis					
Immune System and Blood:					

E.g. Systemic Allergic Reaction

Loco-motor System:

E.g. Arthritis

Skin:

E.g. Contact Dermatitis

Head, Eye, ENT:

E.g. Hearing loss

b. Mortality

Death

2. Psychological Harm

E.g. Anxiety

3. Socio-Economic Harm

E.g. Job loss

Results of the Review

The review yielded empirical data on the risks of the 17 included interventions. The searches retrieved from about 3 to 35 references per intervention that the clinical experts/researchers judged to be of sufficient quality to support risk determinations. The number of observations in the individual datasets ranged from 50 to several 100,000. For several interventions, including ECG and EEG, the search identified case reports only.

The reviewed literature used a wide range of approaches to evaluate risks. The two dominant approaches were 1) retrospective analysis of existing data, for example from existing medical or research records, and 2) surveys that ask members of professional societies to report on the incidence of adverse events for a given procedure and time interval. Other methods included prospective studies, case reports, review papers, and published expert opinion. The available data often varied in the types of potential harms that were reported for each intervention, and how these potential harms were defined. Some of the reported likelihoods for a given potential harm also varied considerably. In some cases, the reported upper likelihood differed from the reported lower likelihood by a factor of 20-30 (online appendix 1 and 2).

Table 4 offers a sample table of the available data on epicutaenous allergy skin testing.

Table 4. Risks of epicutaneous allergy skin testing (sample table).

Allergy Skin Test (Epicutaneous): Risks and Complications

Harm	Likelihood (per 100,000)	95% CI (per 100,000)	N	Study Context	Ref.
1. Physical Harm					
a. Morbidity					
Neurological System:					
Pain, similar to or less than pain from blood draw	~100,000	N/A	N/A	Expert opinion	N/A
Immune System and Blood:					
Systemic allergic reaction, involving pruritus, flushing, light-headedness, urticaria, angioedema, asthmatic or hay fever symptoms	15-23	4-55, 9-67	13,218	Patients at the Mayo Clinic (1992-97) with on average 27 pricks per person; limited to <i>respiratory</i> allergy tests; includes children	(12)
Systemic allergic reaction	19	5-70	10,400	Patients in private practice (1976-1989); includes both epicutaneous and intradermal allergy tests; includes children	(13)
Systemic allergic reaction	20	7-57	16,204	Epicutaneous tests (all kinds of allergies) in healthy volunteers (1976-1980)	(14)
Systemic allergic reaction	30	13-71	16,505	Patients at the Mayo Clinic (1992-97) with on average 27 pricks per person; all kinds of allergy tested; includes children	(12)
Systematic allergic reaction	351	137-900	1,138	Patients with <i>diagnosed food allergy</i> undergoing <i>food allergy skin tests</i> (on average 30 pricks per person); includes children	(15)
Systemic allergic reaction	521	239-1,132	1,152	Percutaneous tests (all kinds) in 1,152 children	(16)
Systemic allergic reaction	6,522	5,236-8,097	1,152	Percutaneous tests (all kinds) in 1,152	(16)

children <6 months					
Mild systemic allergic reaction, involving self-limiting hay fever symptoms, urticaria requiring antihistamine treatment	~13-26	N/A	N/A	Expert opinion	N/A
Moderate to severe systemic allergic reaction, involving asthmatic symptoms and drop in blood pressure, requiring subcutaneous epinephrine treatment ± intubation	~2-5	N/A	N/A	Expert opinion	N/A
Skin:					
Local allergic reaction with intensive, transient pruritus (5-15 min)	~50,000	N/A	N/A	Expert opinion	N/A
b. Mortality					
Death	0	0-24	16,204	Percutaneous tests (all kinds) in 16,204 healthy volunteers (1976-1980)	(14)
Death	0	0-23	16,505	Patients evaluated for allergic diseases at the Mayo Clinic (1992-97); includes children	(12)
Death	0	0-36	10,400	Patients in private practice (1976-1989); includes both epicutaneous and intradermal allergy tests; includes children	(13)
Death	0	N/A	N/A	Fatalities as reported by American Academy of Allergy and Immunology (AAAAI) members; all skin tests (<i>not only epicutaneous</i>) performed between 1985-89	(17)
Death	Case report	N/A	1	Fatalities as reported by AAAAI members (response rate 25%) representing 646 institutions; all epicutaneous skin tests performed between 1990-2001; death from anaphylactic shock in asthmatic patient	(18)

				undergoing 90 food prick tests (implausibly high number of tested allergens); “To our knowledge, this is the first report of a fatality associated exclusively with epicutaneous testing”	
Death	Case report	N/A	6	Fatalities as reported by American Academy of Allergy and Immunology (AAAAI) members; all skin tests (<i>not only epicutaneous</i>) performed between 1959-84	(19)

Quality of data

- Retrospective data analysis (12, 13), physician surveys (17-19), prospective study (13)
- Review articles (20, 21)
- Data include pediatric patients without specifying risks for children; only specific pediatric data (16)
- Incidence of adverse events is probably too high: studies do not have placebo control arms, so it is possible that random allergic reaction coincided with allergy skin testing (evidence for this in (12))

Background information

- *Each epicutaneous skin test involves on average 15-30 pricks per patient (testing the same number of allergens)*
- Pain from superficial skin prick is less than pain from blood draw
- Local reaction involves intense pruritus for about 10-15 min
- Systemic reaction
 - o ~85% mild reaction: self-limiting hay fever symptoms involving sneezing, runny nose and itching eyes for ~1h, urticaria with recovery under antihistamine treatment in ≤ 1 h
 - o ~15% moderate to severe reaction: asthmatic symptoms, hypotension, requires epinephrine treatment
- Anaphylactic shock: essentially no risk due to monitoring and early intervention
 - o For comparison: risk of anaphylactic shock in immunotherapy is 1 per 2-2.5 million injections (17)
- ~85% of research focuses on inhaled allergens (asthma studies), ~15% focus on food allergens

Do Existing Data Meet the 4 Criteria for a Research Risk Repository?

While our search was not relevant to determining whether the obtained data would be regarded as trustworthy by relevant stakeholders (*Criterion #1*), it does provide evidence on the extent to which the available data meet the other three criteria (see online appendix 1 for evidence supporting the examples used in this section).

Criterion #2: Are the available data sufficiently robust and up to date? Based on expert opinion, we included only those research interventions in the search that were expected to yield sufficient risk data. Even so, the quality of the available risk data varied greatly. We found essentially no data on interventions generally perceived to be low risk. For example, our search identified only three studies on the risks of i.v. glucose tolerance testing, and one paper that cites some data as a personal communication. While this finding may reflect the fact that serious harms from glucose tolerance testing are unlikely to occur, the repository needs to clearly distinguish an absence of systematic data on risks from systematic data pointing to an absence of risks for a given intervention.

In addition, the data often were based on low numbers of observations, which is also reflected in large confidence intervals. This shortcoming likely will be difficult to overcome by summarizing data from existing datasets due to a surprising lack of common definitions for potential harms or adverse events. Some datasets did not provide any definitions of the reported adverse events. In other cases, the different datasets used conflicting definitions. For example, i.v. catheters pose a risk of phlebitis. But, some studies defined phlebitis as the occurrence of *any* local symptoms, such as redness,

swelling, local heat, tenderness, or a palpable venous cord. Other studies defined phlebitis as requiring the presence of two, sometimes three, of these symptoms.

Finally, some of the available risk data reflect outdated practices. For example, the largest available datasets on the risks of percutaneous liver biopsy were collected in the 1970s/80s when the biopsies were done without ultrasound guidance, typically by general internists without specialist training.

Criterion #3: Do the existing data include the types and magnitudes of all potential harms? The available data focus primarily on the physical risks of the different interventions. Data on psychological risks are scarce, and data on any social or economic risks are virtually non-existent. Similarly, data on potential negligible harms, such as the dizziness that most people experience when undergoing an i.v. glucose tolerance test, are rarely recorded and their likelihood had to be estimated by expert opinion. This makes it difficult to evaluate the magnitude and likelihood of these harms, important steps in evaluating research risks systematically (42).

Criterion #4: Do the available data cover contextual factors that influence risks?

Information on contextual factors – how the intervention is performed, who undergoes it, and who performs it – is often incomplete. For example, although the risks of sedation are significantly influenced by underlying disease, the available data rarely include subset analyses based on the American Society of Anesthesiology (ASA) physical status classification. In addition, there are almost no data on the risks of common research interventions in relevant subpopulations, such as children and pregnant women. For

example, moderate sedation was the only intervention for which we could develop a summary table specifically for children. This limitation precludes the possibility of making useful estimates about the variance of risks across research sites. Such estimates are helpful for local IRBs only to the extent that they are linked to contextual information about local practices and specific populations.

The available data were almost exclusively collected in the clinical context. Because risk data specifically from the research context are largely absent, it is difficult to assess to what extent this finding confounds the available evidence. Recorded risks might be lower in the research setting because research institutions often have more staff, are better equipped, and typically have a higher level of safety monitoring. In addition, the clinical context can introduce confounding factors that might increase risks. For example, the available data on the risks of lymph node biopsy were collected primarily in oncology patients who underwent sentinel lymph node biopsies. These biopsies tend to be deeper than research biopsies and thus run a higher risk of nerve injury than they would in healthy research subjects.

IMPLICATIONS

The Need to Collect More Data

The conclusion that existing risk data are not sufficiently robust, inclusive, and comprehensive raises additional concern that current evaluations of research risks, even when they appeal to the available data, may not accurately reflect the risks actually faced by research participants. This finding provides compelling support for establishing

a Research Risk Repository. Our literature search also offers several lessons pertinent to collecting the data necessary to establish such a repository.

First, significant additional data are needed on a number of common research interventions. To adequately reflect current practice, these data should be collected in multiple institutions and settings. Second, common definitions of adverse events, as well as coding systems to classify different practices and settings, need to be developed for most research interventions. Third, a Research Risk Repository is likely to be limited to interventions for which standard practices are recognized. It is impossible to collate data when there are no standard practices, or when standard practice varies widely, as currently is the case with moderate sedation and several other interventions (online appendixes).

Present Uses of the Available Risk Data

Creating a Research Risk Repository is likely to take a significant amount of time. Given the pressing need to improve risk judgments, reviewers can *cautiously* consider the results of our preliminary literature search when they assess the risks of research interventions. When doing so, IRBs need to carefully judge both the quality of the data and their relevance for the study under review. Notably, the strength of the data varies for the different interventions from very weak to very strong. For example, the literature on gadolinium-based contrast agents, which are used during imaging interventions, includes datasets with hundreds of thousands of observations. Recognizing that these data are lacking in detail and contextual factors, they are overall robust, and thus merit serious consideration (online appendixes). When estimates about the likelihood of particular

harms vary, IRBs should carefully consider the confidence interval for the likelihood point estimates, generally erring on the side of caution and giving more weight to higher likelihood estimates. IRBs also need to judge, based on the limited information available on contextual factors, to what extent the available data apply to local conditions. When the institutional setting and practices or the study population differ significantly from the existing evidence, IRBs should ask investigators to provide their own data.

ESTABLISHING AND MAINTAINING A RESEARCH RISK REPOSITORY

Creation of a Research Risk Repository should proceed in several steps. First, an appropriate source of funding will have to be identified. To promote the trust of all stakeholders in research, funding should probably be public (e.g., from NIH and/or the FDA). Second, an interdisciplinary committee of experts that oversees the repository will be needed. The committee could be charged with making initial decisions about the repository, in particular which research interventions to focus on first. Third, interdisciplinary groups will be needed to systematically review existing data on the identified interventions, using established methods, such as those developed by the Cochrane Collaboration (43). This effort could be extended by a call for investigators to submit unpublished data.

Fourth, based on this work, the committee of experts should develop common definitions of adverse events and potential harms where necessary. It would also be useful to define standard practices for each intervention and develop codes for these practices, as well as codes for relevant characteristics of research institutions (e.g., number of staff, training institution). Fifth, in parallel to these efforts, a team of

statisticians and database managers could develop the basic structure and functions of the database. This team could start by collating and analyzing the published data with those newly submitted by investigators, drawing on established methods for metaanalysis, while preparing a mechanism for ongoing data collection and analysis.

Data could be transferred from the research institutions to a centralized database by means of a computer program that allows investigators or other research staff to directly export their data to the Research Risk Repository, limiting the effort needed specifically to complete this task. Sixth, it might make sense to start data collection at large centers, such as NIH or the Mayo Clinic, and to evaluate the experience there before progressively expanding the effort. Finally, to ensure the Research Risk Repository provides up to date information, it should be maintained and updated on a regular basis.

LIMITATIONS AND CHALLENGES

An inherent limitation of a Research Risk Repository is that it requires updating as novel research interventions emerge. Creation and maintenance of a Research Risk Repository will also be laborious and costly. However, the costs and effort of establishing a repository would probably not exceed the range for a larger research project with a budget of several million dollars. The continuous costs of maintenance and administration should be dramatically lower. While the total costs of a Research Risk Repository would be significant, they should be evaluated within the larger context of public research expenditures of more than \$30 billion per year in the U.S. alone (44). Several million dollars seem to be a reasonable investment to protect research

participants from excessive risks, while promoting important research consistent with adequate subject protection.

Another challenge is to achieve trust in the Research Risk Repository by the various stakeholders in research. To achieve this, it is important that the repository be operated by an independent non-profit organization, overseen by a committee of independent experts, and made publicly available (e.g., as an online version that people could use to research the risks of common research interventions). Creation of a low-threshold mechanism for complaints about the repository should also help to foster trust.

CONCLUSION

IRBs are charged with evaluating the risks of research interventions. To ensure that participants are protected from excessive risks, and acceptable research is allowed to proceed, it is vital that these judgments reflect the actual risks posed by research interventions. However, IRBs vary greatly in how they classify the risks of research interventions. This variation traces, at least in part, to the fact that few systematic data are available on the risks of research interventions; IRBs consequently evaluate research risks without access to the relevant empirical data.

The present analysis suggests that establishing a Research Risk Repository – a centralized database that maintains systematic data on the outcomes of research interventions – would offer an effective solution to this concern. A preliminary literature search suggests that existing risk data are insufficient for creating such a repository, revealing the need for mechanisms to systematically collect data on the risks of research interventions. In the interim, IRBs might cautiously consider the existing data collated

here when evaluating research risks, provided they are cognizant of their many limitations.

Acknowledgements

This work was completed as part of the authors' official duties as employees of the NIH Clinical Center. However, the NIH had no role in the preparation or review of the manuscript. Thanks to the many experts in biomedical research and research ethics who gave invaluable input on this project; to Ari Hoffman, Elizabeth Gilliam-Hertz, and Leanne Stunkel for research assistance; to Karen Smith for supporting the literature search; and to Ulrike Held for statistical advice and support. Annette Rid was supported by the Swiss National Science Foundation.

Conflict of interest statement

The authors have no conflicts of interest related to this work.

Disclaimer

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The opinions expressed are the authors' own. They do not represent any position or policy of the National Institutes of Health, Public Health Service, or Department of Health and Human Services.

REFERENCES

1. Slovic P. Perception of risk. *Science* 1987; 236: 280-285.
2. Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. *Science* 1974; 185: 1124-1131.
3. Weinstein N. Optimistic biases about personal risks. *Science* 1989; 246: 1232-1233.
4. Slovic P. The perception of risk. *Risk, society, and policy series* 2000: xxxvii, 473 p.
5. Shah S, Whittle A, Wilfond B, et al. How do institutional review boards apply the federal risk and benefit standards for pediatric research? *Jama* 2004; 291: 476-482.
6. Lenk C, Radenbach K, Dahl M, et al. Non-therapeutic research with minors: how do chairpersons of German research ethics committees decide? *J Med Ethics* 2004; 30: 85-87.
7. van Luijn HE, Aaronson NK, Keus RB, et al. The evaluation of the risks and benefits of phase II cancer clinical trials by institutional review board (IRB) members: a case study. *J Med Ethics* 2006; 32: 170-176.
8. Levine RJ. *Ethics and Regulation of Clinical Research*: Urban & Schwarzenberg; 1986.
9. Boulton F. Evidence-based criteria for the care and selection of blood donors, with some comments on the relationship to blood supply, and emphasis on the management of donation-induced iron depletion. *Transfus Med* 2008; 18: 13-27.
10. Brenner DJ, Doll R, Goodhead DT, et al. Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. *Proc Natl Acad Sci U S A* 2003; 100: 13761-13766.
11. Wilson EB. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* 1927; 22.
12. Valyasevi MA, Maddox DE, Li JT. Systemic reactions to allergy skin tests. *Ann Allergy Asthma Immunol* 1999; 83: 132-136.
13. Lin MS, Tanner E, Lynn J, et al. Nonfatal systemic allergic reactions induced by skin testing and immunotherapy. *Ann Allergy* 1993; 71: 557-562.
14. Turkeltaub PC, Gergen PJ. The risk of adverse reactions from percutaneous prick-puncture allergen skin testing, venipuncture, and body measurements: data from the second National Health and Nutrition Examination Survey 1976-80 (NHANES II). *J Allergy Clin Immunol* 1989; 84: 886-890.
15. Codreanu F, Moneret-Vautrin DA, Morisset M, et al. The risk of systemic reactions to skin prick-tests using food allergens: CICBAA data and literature review. *Eur Ann Allergy Clin Immunol* 2006; 38: 52-54.
16. Devenney I, Falth-Magnusson K. Skin prick tests may give generalized allergic reactions in infants. *Ann Allergy Asthma Immunol* 2000; 85: 457-460.
17. Reid MJ, Lockey RF, Turkeltaub PC, et al. Survey of fatalities from skin testing and immunotherapy 1985-1989. *J Allergy Clin Immunol* 1993; 92: 6-15.
18. Bernstein DI, Wanner M, Borish L, et al. Twelve-year survey of fatal reactions to allergen injections and skin testing: 1990-2001. *J Allergy Clin Immunol* 2004; 113: 1129-1136.
19. Lockey RF, Benedict LM, Turkeltaub PC, et al. Fatalities from immunotherapy (IT) and skin testing (ST). *J Allergy Clin Immunol* 1987; 79: 660-677.
20. Lockey RF. Adverse reactions associated with skin testing and immunotherapy. *Allergy Proc* 1995; 16: 293-296.
21. Bernstein IL, Li JT, Bernstein DI, et al. Allergy diagnostic testing: an updated practice parameter. *Ann Allergy Asthma Immunol* 2008; 100: S1-148.

22. Despotis GJ, Goodnough LT, Dynis M, et al. Adverse events in platelet apheresis donors: A multivariate analysis in a hospital-based program. *Vox Sang* 1999; 77: 24-32.
23. McLeod BC, Price TH, Owen H, et al. Frequency of immediate adverse effects associated with apheresis donation. *Transfusion* 1998; 38: 938-943.
24. Laspina SJ, Browne MA, McSweeney EN, et al. QTc prolongation in apheresis platelet donors. *Transfusion* 2002; 42: 899-903.
25. Bolan CD, Greer SE, Cecco SA, et al. Comprehensive analysis of citrate effects during plateletpheresis in normal donors. *Transfusion* 2001; 41: 1165-1171.
26. Bolan CD, Wesley RA, Yau YY, et al. Randomized placebo-controlled study of oral calcium carbonate administration in plateletpheresis: I. Associations with donor symptoms. *Transfusion* 2003; 43: 1403-1413.
27. Bolan CD, Yau YY, Cullis HC, et al. Pediatric large-volume leukapheresis: a single institution experience with heparin versus citrate-based anticoagulant regimens. *Transfusion* 2004; 44: 229-238.
28. Leitman SF, Boltansky H, Alter HJ, et al. Allergic reactions in healthy plateletpheresis donors caused by sensitization to ethylene oxide gas. *N Engl J Med* 1986; 315: 1192-1196.
29. Dutcher JP, Aisner J, Hogge DE, et al. Donor reaction to hydroxyethyl starch during granulocytapheresis. *Transfusion* 1984; 24: 66-67.
30. Lazarus EF, Browning J, Norman J, et al. Sustained decreases in platelet count associated with multiple, regular plateletpheresis donations. *Transfusion* 2001; 41: 756-761.
31. Lewis SL, Kutvirt SG, Bonner PN, et al. Effect of long-term platelet donation on lymphocyte subsets and plasma protein concentrations. *Transfus Sci* 1997; 18: 205-213.
32. Winters JL. Complications of donor apheresis. *J Clin Apher* 2006; 21: 132-141.
33. Popovsky MA, McCarthy S, Hawkins RE. Pseudoaneurysm of the brachial artery: a rare complication of blood donation. *Transfusion* 1994; 34: 253-254.
34. Dettke M, Buchta C, Bieglmayer C, et al. Short and long term effects of citrate on bone metabolism and bone mineral density in healthy plateletpheresis donors. *J Clin Apheresis* 2003; 18: 87.
35. Kose S, Mandiracioglu A. Fear of blood/injection in healthy and unhealthy adults admitted to a teaching hospital. *Int J Clin Pract* 2007; 61: 453-457.
36. Bienvenu OJ, Eaton WW. The epidemiology of blood-injection-injury phobia. *Psychol Med* 1998; 28: 1129-1136.
37. Marks I. Blood-injury phobia: a review. *Am J Psychiatry* 1988; 145: 1207-1213.
38. Agras S, Sylvester D, Oliveau D. The epidemiology of common fears and phobia. *Compr Psychiatry* 1969; 10: 151-156.
39. Hamilton JG. Needle phobia: a neglected diagnosis. *J Fam Pract* 1995; 41: 169-175.
40. Basler HD, Steinfeld C, Hartnack U, et al. The safety of and the strain on the donor during thrombocytapheresis on the cell separator. A psychological study. *Infusionsther Klin Ernahr* 1987; 14 Suppl 4: 36-42.
41. Ringwald J, Lange N, Rabe C, et al. Why do some apheresis donors donate blood just once? *Vox Sang* 2007; 93: 354-362.
42. Rid A, Emanuel EJ, Wendler D. Evaluating the risks of clinical research. *Jama* 2010; 304: 1472-1479.
43. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0, March 2011)2011: Available from: <http://www.cochrane.org/training/cochrane-handbook>.

44. National Institutes of Health. NIH budget2011 24 April, 2011]: Available from: <http://www.nih.gov/about/budget.htm>.